

# 1 Overview and Summary

## 1.1 Executive Summary

The Houston-Galveston-Beaumont (HGB) ozone non-attainment region encompasses a large and diverse emission inventory (EI) and features complex coastal meteorology. As a result, this region has proven difficult to simulate with air quality models combining meteorology and chemistry. Over the last eight years, considerable resources have been spent on surface and aloft measurements and on development of inputs for models of HGB. These intense efforts have produced a simulation for August and September of 2000 that was used as the basis of regulations adopted in 2004. Although this simulation is only a snapshot of one summer in HGB, it has become a “map” for how ozone is formed there. This map has guided policy makers to implement precursor reductions so that the region can comply with the federal one-hour ozone standard. With the introduction of a new federal eight-hour ozone standard, and the likely implementation of yet a lower eight-hour ozone standard, there is now a need to verify the last round of control policies or to formulate new control policies for HGB. In response to this need, resources have been devoted to produce a preliminary simulation for the summer of 2006. This 2006 simulation is a new snapshot of how ozone is formed in HGB. Are the features for ozone production found in the new 2006 simulation similar to those in the 2000 simulation? In other words, if the 2000 simulation depicts a few days that are “representative” of ozone formation processes that needed controls, are the ozone formation processes in the 2006 simulation comparable and will they respond to emissions reductions in the same way as those in the 2000 simulation, or will new types of controls strategies be needed for 2006 and the eight-hour ozone standard?

This work begins to explore the similarities and differences of ozone formation represented by these two important simulations. Understanding these cannot be done by merely contrasting the predicted ozone concentration fields (i.e., plots of 8-h ozone), but requires a deeper investigation into how the models simulate the various physical and chemical processes that are combined in the model to produce the

concentration predictions. The processes are represented not only by the formulations selected in configuring the models, but also by various inputs supplied to the model for each simulation. By quantifying the process magnitudes and the competition among the physical and chemical processes in the model simulation, we can gain a fundamental understanding of how the model produces its ozone (and other relevant species) concentrations. The method of quantification and explaining used herein is called Process Analysis and it was pioneered by UNC in the early 1990's [11]. The method has two components, one internal to the photochemical model that records the magnitudes of each of the processes in the model and one external that post-processes these recorded magnitudes into meaningful process parameters and provides visualization. Both CAMx, the model used by TCEQ to simulate the 2000 episode, and CMAQ, the model used by the University of Houston for the 2006 episode, have the "internal components" to output the Process Analysis model data. The UNC team has reformulated the original post-processing codes and has incrementally improved them in previous TERC/HARC sponsored work; these are described in the H60 project report [29].

It may seem to the reader that comparing two different episodes simulated by two different models is not very meaningful. If all we were comparing were the predicted ozone concentration fields, this would have merit. By quantifying and comparing the most fundamental properties and process magnitudes, however, such comparisons across models for the same episodes and across episodes for different model becomes not only meaningful, it reveals issues and provides insights not obtainable by any other means. Both models have to calculate the same physical and chemical processes that produce ozone. If there are substantial differences in these physical and chemical processes across episodes, even if the predicted ozone is similar, then the system's response to simulated emission controls might be quite different, perhaps being responsive to NO<sub>x</sub> reductions in one case and to VOC reductions in the other.

We always begin by quantifying the "reactor volume" that is loaded with initial concentrations and into which the emission are diluted. How do these volumes differ between the simulations? Next, we quantify the emissions in each scenario and the effect these have, by themselves, on the concentrations in the reactor volume. How do these differ between the simulations? Finally, we quantify the chemical parameters of radical initiation, radical propagation, radical termination and the closely coupled source of NO, its oxidation to NO<sub>2</sub> and its photolysis to recreate the NO and make O<sub>3</sub>, and its loss to non-photolytic nitrogen species. Which VOCs were reacted with the ·OH radical and how much NO was oxidized by the organic radicals created; what secondary aldehydes were created? This is a fundamental chemical engineering approach to reactor analysis: mass balance and reaction rate factors. These questions have meaningful answers regardless of which episode and which model was used to

create the simulation.

There are two other important considerations that make comparing models and episodes relevant:

- Keith Beven’s [1] “equifinality” concept—that for mechanistic modeling of complex environmental systems (which are after all are clearly “open systems”) there are many different model structures and many different parameter sets within a chosen model structure that may be acceptable in reproducing the observed behavior of that system. Which of these acceptable ‘fits’ are more likely?
- Roth, Reynolds, and Tesche’s [22] “compensating errors” concept—one wrong process rate off-setting another wrong process rate and thus the model gives, in Jack Calvert’s words, “seemingly meaningful results.” Are there likely compensating errors in these simulations?

These are much more complicated and complex questions and we will not be able to do more than hint at their answers in this brief study, but our results can focus attention on what evidence and what approaches might lead to meaningful answers.

Because of lack of time resulting from the delay due to difficulties with the meteorological simulations, we had to limit our comparisons of both episodes to one high ozone day with similar meteorological phenomena. We were able to find a similar type of ozone formation that occurred in both 2000 and in 2006: an East to West wind field that carried pollutants from the ship channel west into downtown Houston. This is certainly a relevant—but not the only—case of interest to policy makers.

Our analysis of the 2000 and 2006 episodes focused on downtown Houston and the Houston Ship Channel on August 25, 2000 and August 17, 2006. From 2000 to 2006 there was less than a 13% difference in the two models’ emissions of ozone precursors in the focus region, so we did not expect a large difference in ozone formation processes. The largest differences between the 2000 case and the 2006 case were in the meteorological simulations (both done with the MM5 model). The 2006 MM5 simulation performed by Dr. Byun’s group at UH resulted in a significantly larger extent of vertical mixing in the 2006 episode than that was predicted in the 2000 scenario.

To make this difference clear, we created visualizations of the volume in each model cell that had extensive mixing among the model’s layers (see for example Figure 2.15). We call the volume rendered in the figure the effective Model Mixing Volume (MMV) for each cell for that hour, that is, the volume into which the surface emissions are diluted in the model. The domain-wide average height of the MMV in

the 2006 simulation was almost double that of the 2000 simulation (see Figure 2.13) and it rose faster in 2006 than in 2000.

To be precise, this MMV is not the same as the so-called ‘pbl’ or planetary boundary layer height measured by the sounders and aircraft, as it has been discretised by the chemical model’s layer structure. In the model, once the pbl height exceeds the top of the current model layer, the well-mixed character of the model’s cells effectively distributes material to the top of the next layer layer, resulting in a kind of “numerical” dilution. Likewise, material that was in the aloft layers of the model can then be “entrained” into the more well-mixed volume below.

Our Process Analysis focused on a central  $6 \times 13$  rectangle of  $4\text{-km}^2$  cells over the center of Houston that we call CHOU. The MMV for this area (an example one hour shown in Figure 2.28) is also the volume used each hour for the aggregated chemical analysis we performed. We describe the formation of  $\text{O}_3$  in this entire volume, so that we are focused on overall formation processes and not on those in ‘hot-spots’ or dominated by specific emission sources.

On August 17, 2006, in the CHOU region, there was an unusual 1000 meter rise in the 2006 simulation’s MMV average height at 0900 LST (see Figure 2.27). At no time during the 2000 simulation did such a rapid and early increase in MMV occur. This is a most distinguishing feature between the two scenarios, more so than the emissions.

To our surprise, the analysis for CHOU—with the second highest ozone predicted in the 2006 model simulation—showed a significantly different chemical condition and ozone cause compared to the same region in the 2000 model: the ship channel and central Houston became  $\text{NO}_x$ -limited in the 2006 model by 1100 CST. In the 2000 episode, the same area, with only 10% more  $\text{NO}_x$  emitted, was  $\text{NO}_x$ -inhibited. Indeed, while the VOC-to- $\text{NO}_x$  ratio in the 2006 emissions was close to 2:1, the predicted average VOC-to- $\text{NO}_x$  ratio based on the hourly model-predicted concentrations in the MMV was close to 9:1. In the 2000 model, this ratio was 4:1. The VOC composition in the 2006 MMV is much more ‘aged’. Put another way, VOCs persist in the MMV much longer than does  $\text{NO}_x$ . This model shows a clear example of how *observed* VOC and  $\text{NO}_x$  may not be easily related to *emitted* VOC and  $\text{NO}_x$  and is a point against using *observed* concentrations to infer emissions, without understanding the age distribution of the observed species.

The results described here are almost entirely due to the differences in the timing and magnitude of the predicted MMV in the CHOU region on this day, i.e., to the meteorological inputs to the model and not very much due to any changes in the emissions inventory for this simulation. It is important to note that the meteorological input files used in the 2000 episode had undergone detailed scrutiny resulting in at least three generations of changes. The 2006 meteorological input file, however, is

nearly a first generation file and is still a preliminary model for meteorology in 2006. Until a much more in-depth evaluation of the many vertical and sounder measurements taken during 2006 has been completed, we do not know if this *simulated* large MMV is a reasonable representation of the actual atmospheric condition in 2006.

With the limited time between receiving model data and the termination of the project we were unable to expand our analysis to other days in the 2006 episode to examine similarities with August 17. Future work will involve the expansion of our analysis data set to include all the 2005 modeling episodes developed by the TCEQ. We will then use a similar analysis methodology to compare ozone formation processes in 2000, 2005, and 2006 episodes. The goal of this analysis is to use these three snapshots to understand the variability in formation processes in HGB and to investigate how (or how not) these respond to likely available controls.

## 1.2 Purpose

In this project the University of North Carolina (UNC) research team evaluated a 4-km resolution CMAQ simulation for the episode period August 14, 2006 to Sept 7, 2006 using version TMNS11n2 from the the University of Houston's (UH) MM5 meteorological simulations for the TexAQS II intensive period, combined with the UH estimated and partially imputed emissions inventory (F2A) for this period. The purpose of the evaluation was (1) to characterize the relative contributions of NO<sub>x</sub> and VOC emissions to high ozone concentrations under these new meteorological conditions, and (2) to use UNC's advanced process analysis techniques (PA) to investigate the effects of various additional radical sources including formaldehyde and nitrous acid (HONO) emissions on the ozone predictions. To perform the latter tasks, UNC had to re-simulated the episode with various changes in the emissions while computing the PA outputs. Comparisons were also made with the TCEQ 2000 modeling episode used for the HGB SIP.

In addition to accomplishing those tasks UNC also made further advances in the Process Analysis system software that allows for four-dimensional plume tracking. The system was also expanded with the ability to compare predicted process rates with a battery of measurements using a constrained steady state simulation using various photochemical reaction mechanisms.

## 1.3 Approach

The tasks outlined above were achieved in two parallel efforts. One UNC team focused on the model performance evaluation, while a second team made the modifications to

the process analysis tools.

## Model Performance Evaluation

UNC was tasked to assess the modeling performance of the CMAQ simulation. This simulation used inputs developed at the University of Houston (UH) for a modeling episode that coincided with the Texas Air Quality Study II (TexAQS) in the summer of 2006. Our evaluation of this simulation included a site-by-site comparison of model predictions using a comprehensive observational dataset, and three modeling sensitivity studies. We also used advanced process analysis tools to quantify several chemical parameters important to ozone formation. As a result of this work we developed a modeling evaluation programming suite that is able to efficiently incorporate the latest observational measurements and has the flexibility to integrate rapidly developing modeling data.

Our model evaluation used the `pyPASS` system to integrate modeling and observational data for a detailed performance evaluation. The capabilities and development of the `pyPASS` system is described in detail in the H60 report [29]. UNC integrated into `pyPASS` all available measurement data collected during the summer of 2006 from the TexAQS II field program and from 83 public and private surface monitor sites. The details of this observational dataset are described in Chapter 2 and metadata and summaries are provided in Appendix A. We also integrated all modeling data created by the University of Houston (UH). From these datasets we created plots detailing the model's ability to predict ambient pollutant concentrations and wind fields.

The modeling data used in this study was not available until November 12, 2007 leaving little time for a detailed analysis of the hundreds of plots generated by the `pyPASS` system. As a preliminary analysis, we avoided modeling days with known meteorological errors and low ozone production and also focused our attention on monitors that are located in the urban and industrial sections of Houston. We then performed analysis of pollutant concentrations and meteorological data; results are detailed in Chapter 2. In addition to this analysis we also created three sensitivity studies and quantified the model's response when emissions of formaldehyde and HONO were increased. The H60 project showed that ozone production in the 2000 Houston simulation was sensitive to increased radical sources. The same process analysis methodology described in the H60 report was also used to quantify the changes in chemical processes in the 2006 simulation. The results of this analysis are also detailed in Chapter 2.

## Process Analysis Development

Anticipating future research needs, such as source apportionment analysis, a second goal for UNC was to advance our process analysis techniques to allow for plume tracking. In addition to this task we also developed techniques that allow the process rates in the air quality model to be compared with observation based constrained steady state (CSS) model process rates for any model chemical reaction mechanism.

### Plume Tracking

We continued the development of Process Analysis to create a protocol and enhance the post-processor programs for Process Analysis that can track a release from its emission source to a receptor site while quantifying the physical and chemical processes of the plume. This required extensive modifications of existing process analysis post processing tools. When PA is enabled, the CMAQ air quality model records the integrated process rate (IPR) and integrated reaction rate (IRR) of every physical process and reaction rate for each model cell, at every timestep [5]. An existing UNC post-processor program, named pyPA, then aggregates the cell IPR and IRR data for a collection of grid cells known as a PA volume. To track a plume from an emission source required that the PA volume be able to move with time within the modeling domain. As part of this project, we expanded the PA post-processor to achieve the spatial flexibility necessary to analyze a moving PA volume. The modified pyPA program uses a 4-D matrix to define the PA volumes hourly 3-D spatial location at the beginning and ending of each simulation hour. This method allows us to quantify the changes in concentration due to PA volume motion and/or expansion. The details of this development and a test application on an industrial VOC release are described in Chapter 3. This method permits us to finally explain precisely the limits on ozone production in such plumes.

### Constrained Steady State Model

A program has been developed that can solve a Differential Algebraic Equation (DAE) system that results from a Constrained Steady State (CSS) model. The CSS model assumption has been used primarily for “observation-based analysis”. The development of this program required using the Python language to “wrap” a FORTRAN subroutine called ‘DAE Solver with Krylov Methods and Rootfinding (DASKR)’ supplied by the Lawrence Livermore National Laboratory. DASKR was chosen since it has been extensively tested for solving stiff DAE systems.

In this project, we developed Python wrapper code for DASKR that will permit CSS modeling to be integrated with pyPA. Large parts of the wrapper code have been

borrowed from “SloppyCell Software” [26]. The wrapper interacts with the FORTRAN subroutine through `f2py` (which is part of NumPy [26]) and can wrap FORTRAN subroutines very efficiently. The wrapper is Pythonic in nature and interacts with the FORTRAN subroutine through an interface file.

Observations of  $O_3$ , NO, CO, Carbon Bond 4 (CB4) speciated hydrocarbons, HCHO,  $H_2O_2$ , and  $SO_2$  from a single monitoring location for a particular time are used in the CSS model to yield the concentrations of rapidly reacting species which are in equilibrium with the set of measured compounds. The rapidly reacting species usually include  $\cdot OH$ ,  $HO_2$ ,  $RO_2$ , and  $NO_2$ . From the computed equilibrium concentrations of rapidly reacting species, we can determine the production rate of  $O_3$  ( $P(O_3)$ ) which can then be compared with that from the model for that monitoring location and time. The details of this development and the test example are described in Chapter 4.

## 1.4 Findings

The finding from this report are based on process based analysis of two high ozone production days: (i) August 25 from the 2000 CAMx simulation used in the TCEQ SIP and (ii) August 17 from the UH 2006 simulation delivered to UNC.

We have already discussed the most important finding in the Executive Summary: the 2006 episode meteorological simulation produce very large Model Mixing Volumes (MMV). The domain wide MMV in the 2006 simulation was almost double that of the 2000 simulation and it rose faster. In the PA focus region, there was a 1000 meter rise at 0900 LST in the 2006 simulation (see Figure 2.27). This rapidly diluted the NO emissions that permitted a more rapid increase in  $O_3$  that increased radicals while the lower  $NO_2$  resulted in less radical termination. Thus, the MMV differences produced a more reactive system in 2006.

While there was a large dilution of  $NO_x$  emissions (and VOC emissions), the lifetime of VOC is much longer than that of  $NO_x$  and the model layers in the morning above the MMV were rich in left-over VOC (but not  $NO_x$ ). Entraining this older VOC into the rapidly diluting mixed layer, resulted in a VOC-to- $NO_x$  ratio shift from 2:1 in the emission to 9:1 in the mixed volume of central Houston. The model became  $NO_x$ -limited by 1100 LST. In the absence of sufficient NO, significant quantities of hydrogen peroxide ( $H_2O_2$ ) were produced in the model from  $HO_2 + HO_2$  termination reactions (see Figure 2.29). In future studies, we will be comparing the model’s predictions with  $H_2O_2$  measurements taken at Moody Tower (which is within the CHOU, see MDYT site in Figure 2.2). The limited time after receiving and analyzing the model data did not permit us to make this comparison in this study.

The 2006 emissions inventory used by UH has some interesting differences com-

pared to the 2000 inventory. Overall, in the 2006 simulation, there was a 21% increase in VOC emissions, mostly due to higher biogenic VOC emissions (see Figure 2.22 and Figure 2.24). The highway system did show a decrease in VOC emissions rates. Overall, in the 2006 simulation, there was 5% less CO and 8% less NO<sub>x</sub> emissions, however, there were increases in NO<sub>x</sub> over the whole domain area with larger decreases in NO<sub>x</sub> over highways and urban areas (see Figure 2.18 and Figure 2.20).

The higher MMV resulted in different concentration distributions and therefore a different chemical outcome. To help understand this, we have prepared tile timeseries plots of the concentrations that would result in each cell from *just* emissions into the volume created by the cell area times the MMV height. These are in ppb. For the case of NO<sub>x</sub> (see Figure 2.19 and Figure 2.21) and VOC (see Figure 2.23 and Figure 2.25) the higher MMV overwhelm the increases in the emissions and resulted in lower (sometimes much lower) emissions-only-generated concentrations in the cells over the total MMV height of each cell.

The difference in MMV heights resulted in significant changes to the emissions and transport of ozone precursors. Table 2.8 shows that much of the VOC species are transported out of PA analysis volume in 2000. In the 2006 episode, however, there are NO<sub>x</sub> and VOC concentration being brought into the PA analysis volume. Both episodes rise in the PA volume height results in dilution of precursor concentrations, but it was nearly twice as much in 2006. With this data we can also calculate the VOC to NO<sub>x</sub> ratios based on emissions and on the average box concentrations. The ratios are similar between both episodes when based on emissions confirming the small changes in the EI from Table 2.7. There is significant difference, however, when based on the average concentrations found in the PA volume. This discrepancy reflects the “carry-over” of VOC and loss of NO<sub>x</sub> from hour to hour and the more stagnant conditions found in 2006. This has implications for observational based assessments or for “inverse modeling” of VOCs.

Although both the 2000 and 2006 base simulation have the same production of new ·OH (37.5 and 37.7 ppb, see Table 2.9 and Table 2.10), the 2006 simulation produced more ozone (133 ppb compared to 116 ppb). The decrease in NO<sub>x</sub> concentrations due to changes in both the MMV and in the 2006 emission inventory, resulted in a dramatic decline in ·OH termination reactions with NO<sub>2</sub>. The result was more ·OH reactions with VOCs and CO rather than with NO<sub>2</sub> a ratio of 6.6 in 2006 compared to only 3.2 in the 2000 scenario. Figure 2.29 clearly shows the decrease in the ·OH + NO<sub>2</sub> reaction pathway in the 2006 episode due to dilution of NO<sub>x</sub>. The figure also shows the amount of hydrogen peroxide made in 2006, and indication of NO<sub>x</sub> limited chemistry. A second indication of a NO<sub>x</sub> limitation is the increase in the NO chain length from 1.58 to 3.77. A third piece of evidence pointing to a NO<sub>x</sub> limited atmosphere is that in 2006 only 65% of HO<sub>2</sub> reacted with NO in the afternoon versus 95% in 2000. Although

both episodes started with the same amount of new  $\cdot\text{OH}$  radicals, the 2006 episode made more ozone mostly because of the dilution of the  $\text{NO}_x$  reducing termination pathways.

Figures 2.29 and 2.30 report the distribution of VOCs that reacted with the  $\cdot\text{OH}$  radical. The simulation in 2006 showed an increase in reactions with  $\text{CH}_4$ , PAR, and secondary species like FORM, ALD2 and isoprene's secondary species ISPD. There was also a decrease in reactions with ETH, OLE, and ISOP. These data suggest that in the 2006 episode, the diluted VOCs resulted in the quick depletion of the highly reactive VOCs and an increase in reactions with the less reactive, yet abundant PAR and secondary products.

Finally, the 2006 simulation was not as responsive to increased sources of new radicals as was the 2000 simulation. This is in part due to much less termination of existing radicals in the 2006 case. This too was due more to the larger MMV heights in 2006 than to reductions in  $\text{NO}_x$  emissions.